MS1: Predictors of Sex-Specific Heat Stress Death

2025-07-15

Variables include:

- Mortality outcome (Died_InTrt)

- Demographic variables (Sex, AgeCategory)

- Baseline blood measures (Glucose, Ketone, Hematocrit)

 # - Treatments: A (Control), B (Heat-Stress)

Load Libraries

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(stringr)
library(tidyr)
library(ggplot2)
library(logistf)
library(lme4)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack

library(car)


## Loading required package: carData


##
## Attaching package: 'car'


## The following object is masked from 'package:dplyr':
##
##     recode

library(ggeffects)
library(tibble)
```

Clear memory, set working directory, read in datafile & define plot theme

```
# Clears memory
rm(list=ls(all = TRUE))

#Set working directory (assumes data file is available in same location as code)
current_path = rstudioapi::getActiveDocumentContext()$path
setwd(dirname(current_path))

#Read Data File
AHDB_Exp1_BloodData <- read.csv("AHDB_Exp1_BloodData.csv")
str(AHDB_Exp1_BloodData)
```

```
## 'data.frame':    256 obs. of  21 variables:
##  $ AgeCategory        : chr  "Older" "Older" "Older" "Older" ...
##  $ ID_Band            : int  4572 4556 4572 4556 4531 4614 4531 4614 4529 4608 ...
##  $ Sex                : chr  "F" "M" "F" "M" ...
##  $ Mom_ID             : int  4112 4287 4112 4287 4115 4340 4115 4340 4111 4124 ...
##  $ Cohort             : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Treatment          : chr  "A" "A" "A" "A" ...
##  $ Died_InTrt         : chr  "No" "No" "No" "No" ...
##  $ TimePoint          : chr  "Baseline" "Baseline" "Treatment" "Treatment" ...
##  $ Glucose_1          : chr  "162" "213" "192" "197" ...
##  $ Glucose_2          : chr  "171" "211" "239" "194" ...
##  $ Glucose_3          : chr  "No" "No" "No" "No" ...
##  $ Ketone_1           : chr  "0.4" "0.2" "0.4" "0.4" ...
##  $ Ketone_2           : chr  "0.3" "0.2" "No" "No" ...
##  $ Ketone_3           : chr  "No" "No" "No" "No" ...
##  $ Plasma_Color       : chr  "U" "No" "U" "U" ...
```

```
## $ Hematocrit               : chr   "56" "50" "56" "56" ...
## $ BleedTime_Min            : chr   "" "" "1:09" "1:09" ...
## $ Glucose_Meter_Lot_Number : chr   "031324A" "022723A" "031324A" "031324A" ...
## $ Ketone_Meter_Lot_Number  : chr   "" "" "" "" ...
## $ Room                     : int   118 114 113 113 118 114 113 113 118 114 ...
## $ Notes                    : chr   "" "" "" "" ...
```

```r
# Define Plot Theme
plot_theme <- theme_classic(base_size = 18) +
  theme(axis.title.y = element_text(vjust = 1.5),
        axis.title.x = element_blank()) +
  theme(plot.margin = unit(c(.3, .3, .6, .6), "cm"),
        line = element_line(linewidth = 1.25)) +
  theme(axis.line.x = element_line(colour = "black", linewidth = 1.25),
        axis.line.y = element_line(colour = "black", linewidth = 1.25)) +
  theme(axis.text.x = element_text(angle = 360, hjust = .5,
                                   margin = unit(c(0.5, 0.5, 0.5, 0.5), "cm"),
                                   color = "black"),
        axis.text.y = element_text(margin = unit(c(0.5, 0.5, 0.5, 0.5), "cm"),
                                   color = "black")) +
  theme(axis.ticks.length = unit(-0.3, "cm"),
        axis.ticks = element_line(color = "black", linewidth = 1.25)) +
  theme(panel.spacing = unit(2, "lines")) +
  theme(strip.background = element_blank()) +
  #theme(legend.position = "none") +
  theme(plot.title = element_blank())
```

Data Processing & Filtering

Filter AHDB_Exp1_BloodData to baseline timepoint only

This is a post hoc analysis of heat-stress induced mortality, therefore treatment A (control) was removed due to individuals not experiencing the heat treatment

```r
AHDB_Exp1_Baseline <- AHDB_Exp1_BloodData %>%
  filter(TimePoint == "Baseline",
         Treatment   != "A")

#Convert Died in Treatment to factor
AHDB_Exp1_Baseline <- AHDB_Exp1_Baseline %>%
  mutate(
    Died_InTrt = factor(Died_InTrt,
                        levels = c("No", "Yes")))



# Define columns to convert to numeric

cols_to_convert <- c("Glucose_1", "Glucose_2", "Glucose_3",
```

```r
                         "Ketone_1", "Ketone_2", "Ketone_3", "Hematocrit")

# Replace "No" with NA in the specified columns
AHDB_Exp1_Baseline[cols_to_convert] <- lapply(AHDB_Exp1_Baseline[cols_to_convert], function(x) {
  x[x == "No"] <- NA
  x
})

# Convert these columns to numeric
AHDB_Exp1_Baseline[cols_to_convert] <- lapply(AHDB_Exp1_Baseline[cols_to_convert], as.numeric)

# Calculate average values for Glucose and Ketones
AHDB_Exp1_Baseline$Glucose_Avg <- rowMeans(AHDB_Exp1_Baseline[, c("Glucose_1", "Glucose_2", "Glucose_3"
                                                    na.rm = TRUE)
AHDB_Exp1_Baseline$Ketone_Avg <- rowMeans(AHDB_Exp1_Baseline[, c("Ketone_1", "Ketone_2", "Ketone_3")],
                                                    na.rm = TRUE)

# Create a Female ZEFI baseline dataset
AHDB_Baseline_Female <- AHDB_Exp1_Baseline %>%
  filter(
    Sex          == "F",
    TimePoint    == "Baseline"
  )

# Create a Younger Female ZEFI baseline dataset
AHDB_Baseline_YoungerFemale <- AHDB_Baseline_Female %>%
  filter(
    Sex          == "F",
    AgeCategory   == "Younger"
  )
```

Data Analysis

Hypothesis 1 . Females are more likely to die from acute heat (~ 43 − 43.5C for 5 hours) than males.(Includes Mom_ID as a random effect)

Calculate Sample Size by Sex (Does not account for age)

```r
AHDB_Exp1_Baseline %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Males   = sum(Sex == "M"),
    Females = sum(Sex == "F")
  )
```

```
##   Total Males Females
## 1    87    19      68
```

Firth Model -> Sex

This runs Firth's logistic regression with the penalized likelihood approach)

Useful for data with reduced sample size and complete separation in variables (DeathInTrt - No Males Died)

```
firth_fit_sex <- logistf(Died_InTrt ~  Sex ,
                        data = AHDB_Exp1_Baseline)
summary(firth_fit_sex)
```

```
## logistf(formula = Died_InTrt ~ Sex, data = AHDB_Exp1_Baseline)
##
## Model fitted by Penalized ML
## Coefficients:
##                  coef  se(coef) lower 0.95  upper 0.95     Chisq          p
## (Intercept) -1.609438 0.3230291  -2.293611 -1.01518560 33.476864 7.211676e-09
## SexM        -2.054124 1.4682063  -6.925818  0.06681891  3.539758 5.991419e-02
##             method
## (Intercept)      2
## SexM             2
##
## Method: 1-Wald, 2-Profile penalized log-likelihood, 3-None
##
## Likelihood ratio test=3.539758 on 1 df, p=0.05991419, n=87
## Wald test = 31.36669 on 1 df, p = 2.136147e-08
```

```
exp(coef(firth_fit_sex))
```

```
## (Intercept)        SexM
##   0.2000000    0.1282051
```

```
exp(confint(firth_fit_sex))
```

```
##                 Lower 95% Upper 95%
## (Intercept) 0.1009014151 0.3623352
## SexM        0.0009820991 1.0691019
```

```
# Sex p = 5.991419e-02 = 0.0599
```

Calculate pseudo-R2 (Proportion of variation of independent variable explained by the dependent variable)

```
ll.null <- firth_fit_sex$loglik[1]
ll.proposed <- firth_fit_sex$loglik[2]
pseudoR2 <- (ll.proposed - ll.null) / ll.null
cat("Pseudo R²:", pseudoR2, "\n")
```

```
## Pseudo R²: 0.05930259
```

```
# Result: 5.9%
```

Predicted probabilities for each Sex

```
death_data <- data.frame(Sex = c("M", "F"))
predicted_prob_sex <- predict(firth_fit_sex, death_data, type = "response")
prob_table <- data.frame(Sex = death_data$Sex, Predicted_Probability = predicted_prob_sex)
print(prob_table)
```

```
##   Sex Predicted_Probability
## 1   M             0.0250000
## 2   F             0.1666667
```
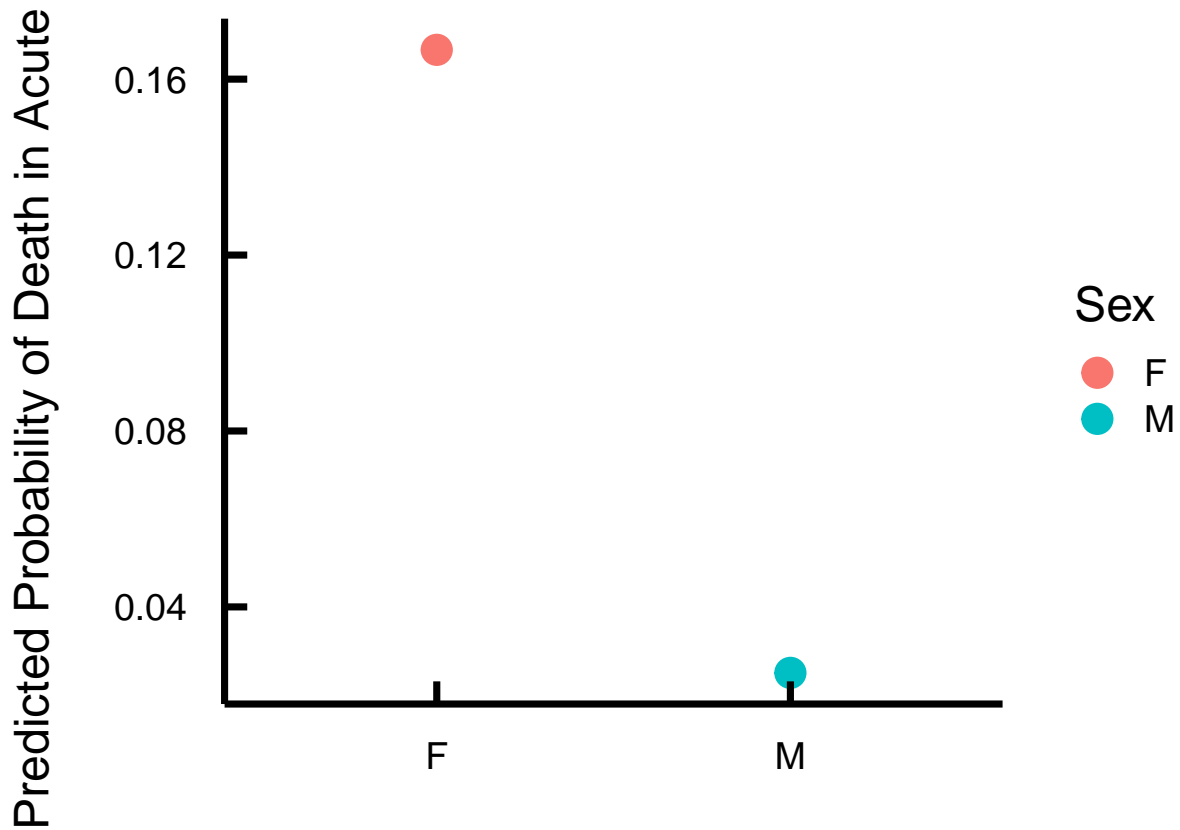
```
# Results: Females have a 16.7 % likelihood of succumbing to acute heat exposure compared to 2.5% for t
```

Plot predicted probabilities

```
predicted_prob_sex_plot <-ggplot(death_data, aes(x = Sex, y = predicted_prob_sex, fill = Sex)) +
  geom_point(aes(color = Sex), size = 5) +
  xlab("Sex") +
  ylab("Predicted Probability of Death in Acute Heat") +
  plot_theme

predicted_prob_sex_plot
```

```
#ggsave("MS1_Predicted_Death.jpg",predicted_prob_sex_plot, dpi    = 250,
#width  = 10, height = 8)
```

Calculate observed death proportions

```
obs_death_sex <- AHDB_Exp1_Baseline %>%
  mutate(Death = if_else(Died_InTrt == "Yes", 1, 0)) %>%
  group_by(Sex) %>%
  summarize(
    N = n(),
    DeathCount = sum(Death, na.rm = TRUE),
    Proportion = DeathCount / N
  )
print(obs_death_sex)
```

```
## # A tibble: 2 x 4
##   Sex       N DeathCount Proportion
##   <chr> <int>      <dbl>      <dbl>
## 1 F        68         11      0.162
## 2 M        19          0      0
```
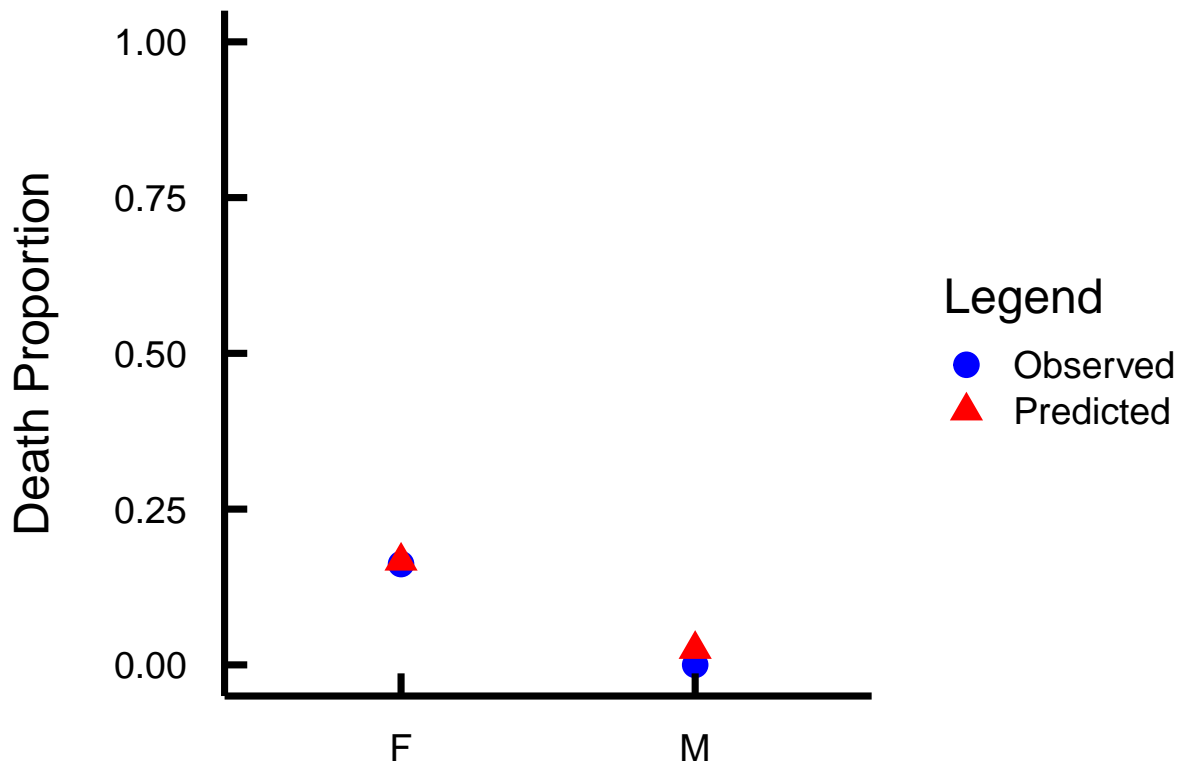
```
# Observed Female Death Proportion: 0.162
# Observed Male Death Proportion: 0
```

Observed vs Predicted Prob of Death Plot

```
pred_obs_sex <- ggplot() +
  # Plot observed death proportions (blue circles)
  geom_point(data = obs_death_sex,
             aes(x = Sex, y = Proportion, color = "Observed"),
             size = 4) +
  # Plot model predictions (red triangles)
  geom_point(data = prob_table,
             aes(x = Sex, y = Predicted_Probability, color = "Predicted"),
             size = 4, shape = 17) +
  scale_color_manual(values = c("Observed" = "blue", "Predicted" = "red"),
                     name = "Legend",
                     labels = c("Observed", "Predicted")) +
  labs(title = "Observed vs. Predicted Death Proportions by Sex",
       x = "Sex",
       y = "Death Proportion") +
  ylim(0, 1) +
  plot_theme

pred_obs_sex
```

ESEB Poster Brynleigh - Observed Mortality by sex and age

```
obs_death <- AHDB_Exp1_Baseline %>%
  mutate(Death = if_else(Died_InTrt == "Yes", 1, 0)) %>%
  group_by(Sex, AgeCategory) %>%
  summarize(
    N = n(),
    DeathCount = sum(Death, na.rm = TRUE),
    Proportion = DeathCount / N
  )
```

```
## `summarise()` has grouped output by 'Sex'. You can override using the `.groups`
## argument.
```

```
print(obs_death)
```

```
## # A tibble: 4 x 5
```

```
## # Groups:   Sex [2]
##   Sex  AgeCategory     N DeathCount Proportion
##   <chr> <chr>       <int>     <dbl>      <dbl>
## 1 F    Older          23         1     0.0435
## 2 F    Younger        45        10     0.222
## 3 M    Older           5         0     0
## 4 M    Younger        14         0     0
```

```r
# Dot Plot
ggplot(obs_death, aes(x = AgeCategory , y = Proportion)) +
  geom_point(aes(color = Sex), size = 5) +  # Points
  labs(
    title = "High Mortality of Young Females During Acute Heat",
    x = "Age Category",
    y = "Observed Death Proportions"
  ) +
  theme_classic()
```
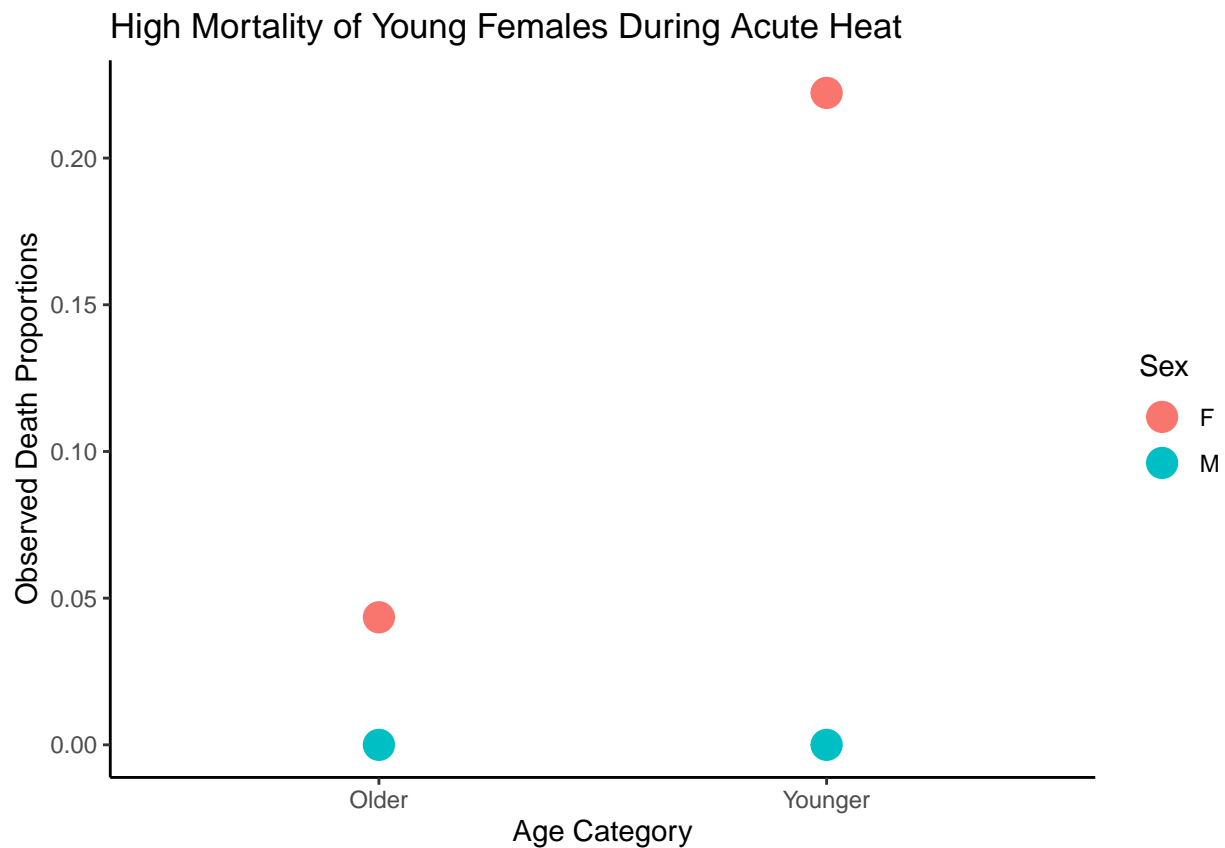


```r
# Bar Graph
# Create a plotting version of the data with a tiny positive value for zero bars
obs_death_plot <- obs_death %>%
  mutate(Proportion_fixed = ifelse(Proportion == 0, 0.001, Proportion))

ggplot(obs_death_plot, aes(x = AgeCategory, y = Proportion_fixed, fill = Sex)) +
  geom_col(position = position_dodge(width = 0.9), width = 0.8) +
```

```
  # Apply custom colors for Sex
  scale_fill_manual(values = c( "#F79646", "#78898F")) +
  labs(
#    title = "High mortality of younger females during acute heat",
    x = "Age Category",
    y = "Observed Mortality"
  ) +
  theme_classic(base_size = 18)
```
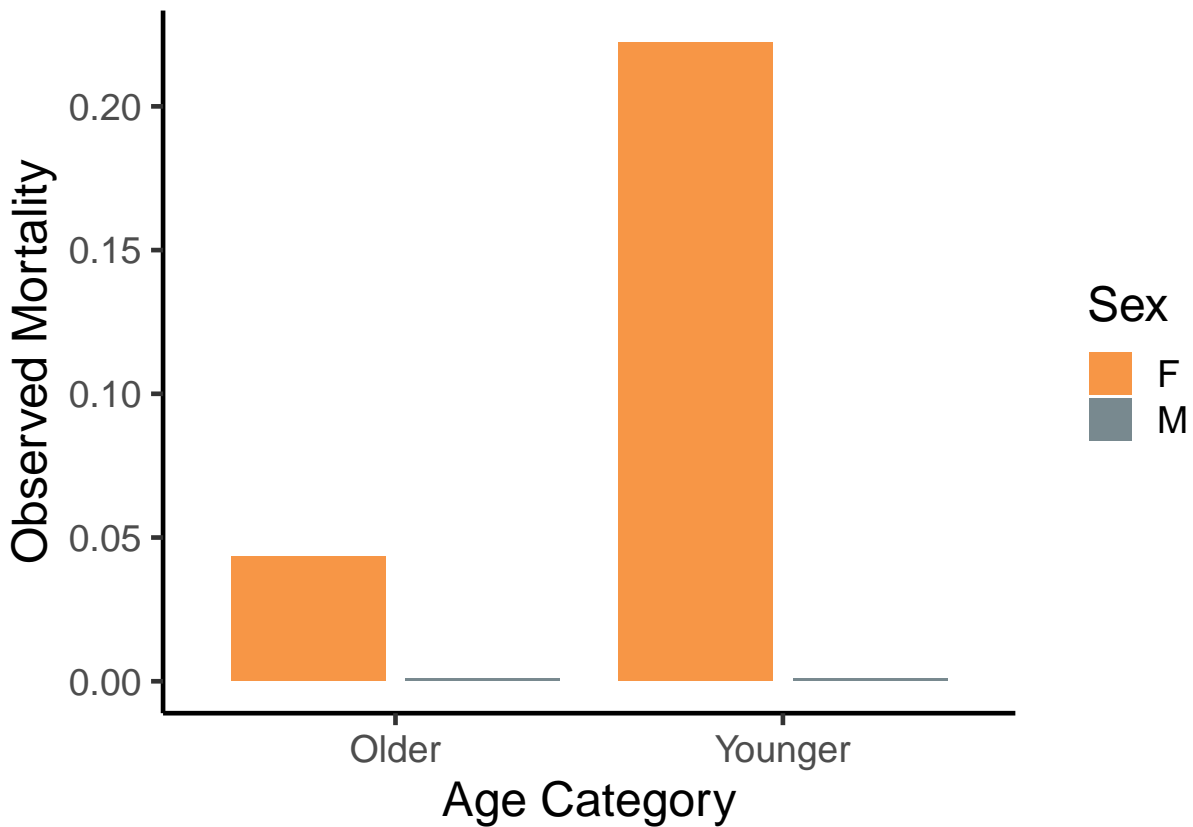

Calculate Sample Size by Age & Sex for Firth Model that includes both sex and age

```
AHDB_Exp1_Baseline %>%
  distinct(ID_Band, Sex, AgeCategory) %>%
  group_by(Sex, AgeCategory) %>%
  summarise(
    Count = n(),
    .groups = "drop"
  )
```

```
## # A tibble: 4 x 3
##   Sex   AgeCategory Count
```

```
##   <chr> <chr>      <int>
## 1 F     Older          23
## 2 F     Younger        45
## 3 M     Older           5
## 4 M     Younger        14
```

Firth Model - Sex + Age

```
firth_fit_sex_age<- logistf(Died_InTrt ~  Sex + AgeCategory, #+ Sex*AgeCategory,
                            data = AHDB_Exp1_Baseline)
summary(firth_fit_sex_age)
```

```
## logistf(formula = Died_InTrt ~ Sex + AgeCategory, data = AHDB_Exp1_Baseline)
##
## Model fitted by Penalized ML
## Coefficients:
##                           coef  se(coef)  lower 0.95 upper 0.95      Chisq
## (Intercept)          -2.705328 0.8325384 -4.89957785 -1.3557985 22.051935
## SexM                 -2.191858 1.4541087 -7.06875834 -0.0486786  4.067841
## AgeCategoryYounger    1.485242 0.9001280 -0.06746041  3.7589476  3.479126
##                                 p method
## (Intercept)          2.653720e-06      2
## SexM                 4.370709e-02      2
## AgeCategoryYounger   6.214754e-02      2
##
## Method: 1-Wald, 2-Profile penalized log-likelihood, 3-None
##
## Likelihood ratio test=7.016814 on 2 df, p=0.02994458, n=87
## Wald test = 28.62498 on 2 df, p = 6.083669e-07
```

```
#Age & Sex interaction was removed by backward regression
```

```
exp(coef(firth_fit_sex_age))
```

```
##       (Intercept)               SexM AgeCategoryYounger
##        0.06684836         0.11170904         4.41603222
```

```
exp(confint(firth_fit_sex_age))
```

```
##                      Lower 95%  Upper 95%
## (Intercept)        0.0074497273  0.2577414
## SexM               0.0008512895  0.9524872
## AgeCategoryYounger 0.9347647226 42.9032520
```

```
# Sex p= 4.370709e-02 = 0.044*
# Age p= 6.214754e-02 = 0.062
```

Calculate pseudo-R2 (Proportion of variation of independent variable explained by the dependent variable)

```
ll.null <- firth_fit_sex_age$loglik[1]
ll.proposed <- firth_fit_sex_age$loglik[2]
pseudoR2 <- (ll.proposed - ll.null) / ll.null
cat("Pseudo R²:", pseudoR2, "\n")
```

```
## Pseudo R²: 0.1266361
```

```
# Pseudo R²: 0.1266361
```

Calculate Predicted Probabilities with both Age & Sex

```
death_data <- data.frame(
  Sex         = c("M","F","F","M"),
  AgeCategory = c("Younger","Older","Younger","Older")
)
predicted_prob_sex_age <- predict(firth_fit_sex_age, death_data, type = "response")
prob_table_sex_age <- data.frame(Sex = death_data$Sex, AgeCategory =death_data$AgeCategory, Predicted_P:
print(prob_table_sex_age)
```

```
##   Sex AgeCategory Predicted_Probability
## 1   M     Younger            0.031924246
## 2   F       Older            0.062659665
## 3   F     Younger            0.227921160
## 4   M       Older            0.007412215
```

```
# Results: Younger Females have a 22.8% likelihood of succumbing to acute heat exposure compared to 3.2
```

Hypothesis 2. Younger adult females are more likely to die from acute heat (~ 43–43.5C for 5 hours) than older young adult females.

#Includes Mom_ID as a random effect???

Firth - Younger Females Only

```
firth_fit_age <- logistf(Died_InTrt ~  AgeCategory,
                         data = AHDB_Baseline_Female)
summary(firth_fit_age)
```

```
## logistf(formula = Died_InTrt ~ AgeCategory, data = AHDB_Baseline_Female)
##
```

```
## Model fitted by Penalized ML
## Coefficients:
##                           coef   se(coef)  lower 0.95  upper 0.95      Chisq
## (Intercept)          -2.708050  0.8432740 -4.90277686   -1.357436  22.049065
## AgeCategoryYounger    1.489893  0.9135197 -0.06377087    3.763601   3.498879
##                               p method
## (Intercept)          2.657691e-06      2
## AgeCategoryYounger   6.141039e-02      2
##
## Method: 1-Wald, 2-Profile penalized log-likelihood, 3-None
##
## Likelihood ratio test=3.498879 on 1 df, p=0.06141039, n=68
## Wald test = 22.33728 on 1 df, p = 2.287235e-06
```

```
exp(coef(firth_fit_age))
```

```
##       (Intercept) AgeCategoryYounger
##        0.06666667         4.43661972
```

```
exp(confint(firth_fit_age))
```

```
##                        Lower 95%   Upper 95%
## (Intercept)          0.007425934   0.2573198
## AgeCategoryYounger   0.938219951  43.1033677
```

```
# Age p= 6.141039e-02 = 0.061
```

Calculate "Pseudo R²

```
ll.null <- firth_fit_age$loglik[1]
ll.proposed <- firth_fit_age$loglik[2]
pseudoR2 <- (ll.proposed - ll.null) / ll.null
cat("Pseudo R²:", pseudoR2, "\n")
```

```
## Pseudo R²: 0.06515706
```

```
# Pseudo R²: 0.06515706
```

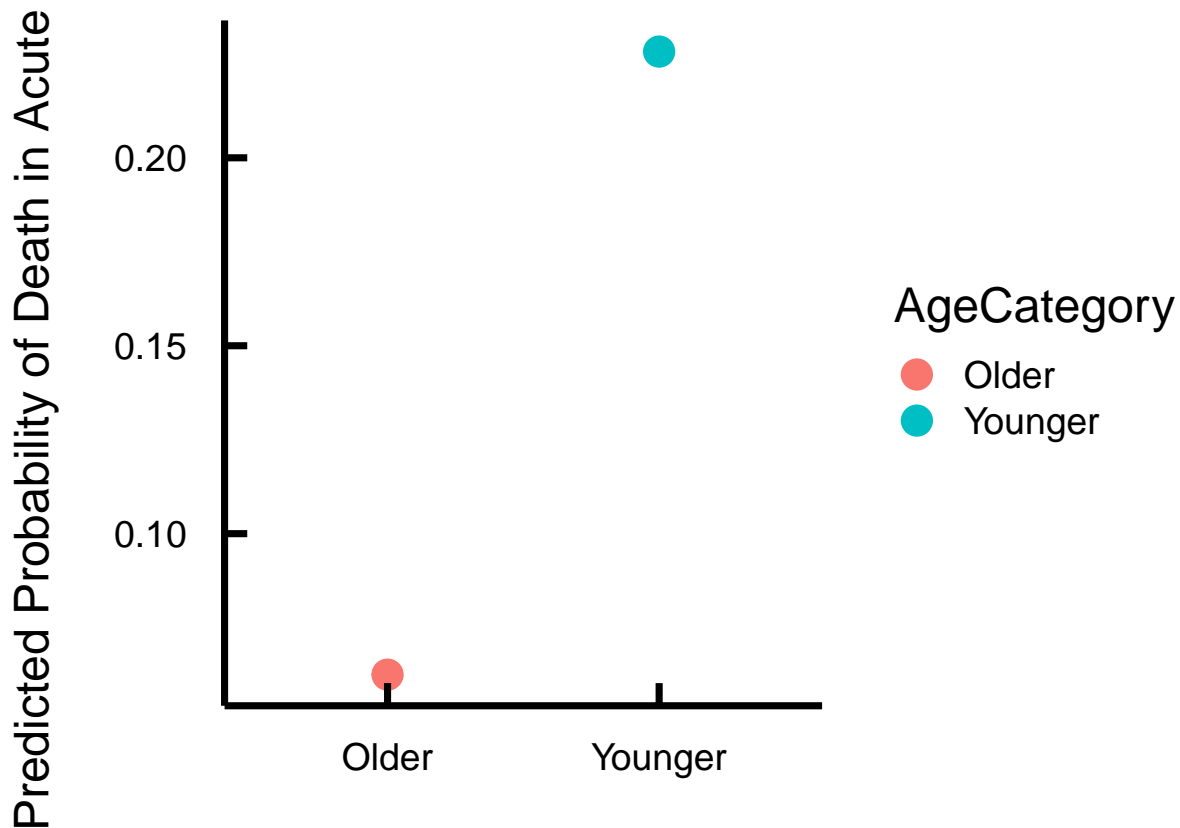# Predicted probabilities for each AgeCategory

```
deathdata_H2 <- data.frame(AgeCategory = c("Older", "Younger"))
predicted_prob_age <- predict(firth_fit_age, newdata = deathdata_H2, type = "response")
prob_table <- data.frame(Age = deathdata_H2$Age, Predicted_Probability = predicted_prob_age)
print(prob_table)
```

```
##        Age Predicted_Probability
## 1    Older             0.0625000
## 2  Younger             0.2282609
```

14

Plot predicted probabilities

```
predicted_prob_age_plot <-ggplot(deathdata_H2, aes(x = AgeCategory, y = predicted_prob_age, fill = AgeC
  geom_point(aes(color = AgeCategory), size = 5) +
  xlab("Age") +
  ylab("Predicted Probability of Death in Acute Heat") +
  plot_theme

predicted_prob_age_plot
```



```
#ggsave("MS1_FemalePredicted_AgeDeath.jpg", predicted_prob_age_plot, dpi    = 250,
#  width  = 10, height = 8)
```

Calculate observed death probabilites for younger females

```
obs_death_age <- AHDB_Baseline_Female %>%
  mutate(Death = if_else(Died_InTrt == "Yes", 1, 0)) %>%
  group_by(AgeCategory) %>%
  summarize(
    N = n(),
    DeathCount = sum(Death, na.rm = TRUE),
```

```
    Proportion = DeathCount / N
  )
print(obs_death_age)
```

```
## # A tibble: 2 x 4
##   AgeCategory     N DeathCount Proportion
##   <chr>       <int>      <dbl>      <dbl>
## 1 Older          23          1     0.0435
## 2 Younger        45         10     0.222
```

Observed vs Predicted Prob of Death Plot

```
obs_death_age <- ggplot() +
  # Plot observed death proportions (blue circles)
  geom_point(data = obs_death_age,
             aes(x = AgeCategory, y = Proportion, color = "Observed"),
             size = 4) +
  # Plot model predictions (red triangles)
  geom_point(data = prob_table,
             aes(x = Age, y = Predicted_Probability, color = "Predicted"),
             size = 4, shape = 17) +
  scale_color_manual(values = c("Observed" = "blue", "Predicted" = "red"),
                     name = "Legend",
                     labels = c("Observed", "Predicted")) +
  labs(title = "Observed vs. Predicted Death Proportions by Age",
       x = "Age",
       y = "Death Proportion") +
  ylim(0, 1) +
  plot_theme

obs_death_age
```
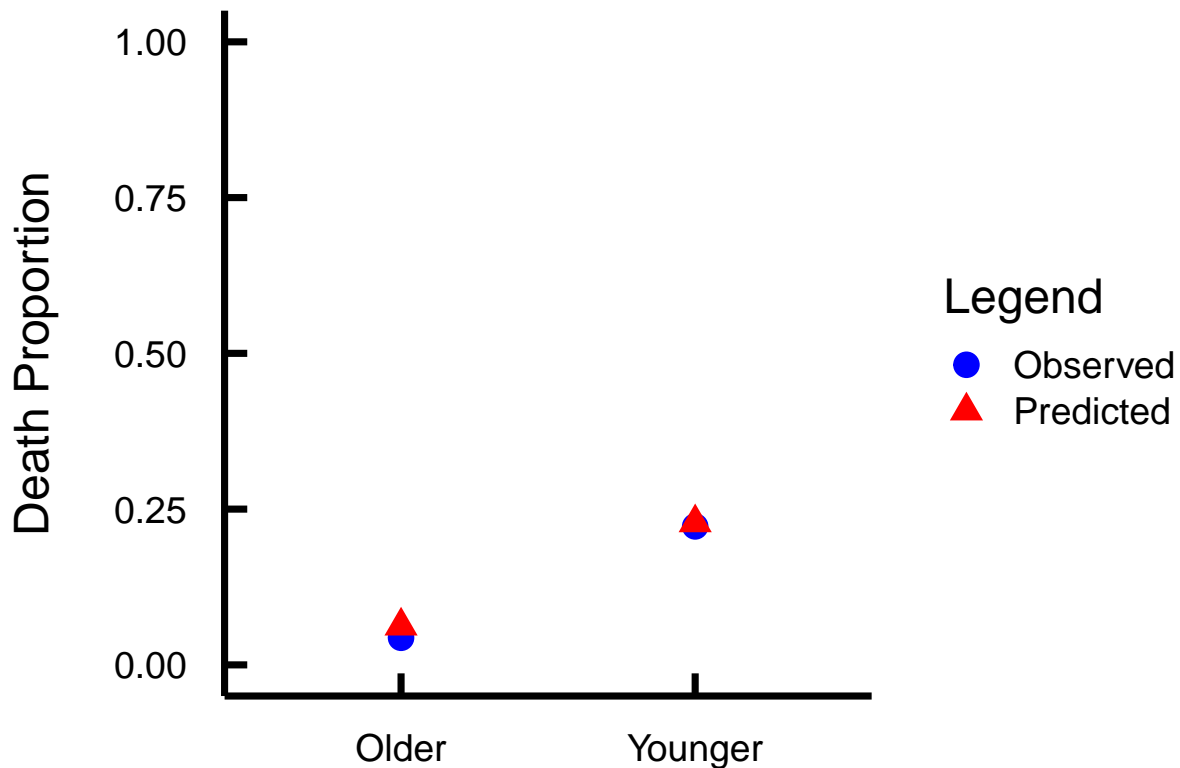
Hypothesis 3. Among young adult females, there will be metabolic or damage variables in the baseline blood samples (2 weeks prior) may correlated with probability of dying due to acute heat (~$43 - 43.5C$ for 5 hours).

#Includes Mom_ID as a random effect

Glucose Model -> To see whether baseline glucose levels predict death in younger females (Includes Individuals with lysed plasma)

Remove rows with missing values

```
Younger_females_Glu_lysed <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Glucose_Avg))
```

Calculate Mean Glucose Values & plot residuals

```
mean(Younger_females_Glu_lysed$Glucose_Avg)
```

```
## [1] 228.9886
```

```
sd(Younger_females_Glu_lysed$Glucose_Avg)
```

```
## [1] 40.25668
```

```
min(Younger_females_Glu_lysed$Glucose_Avg)
```
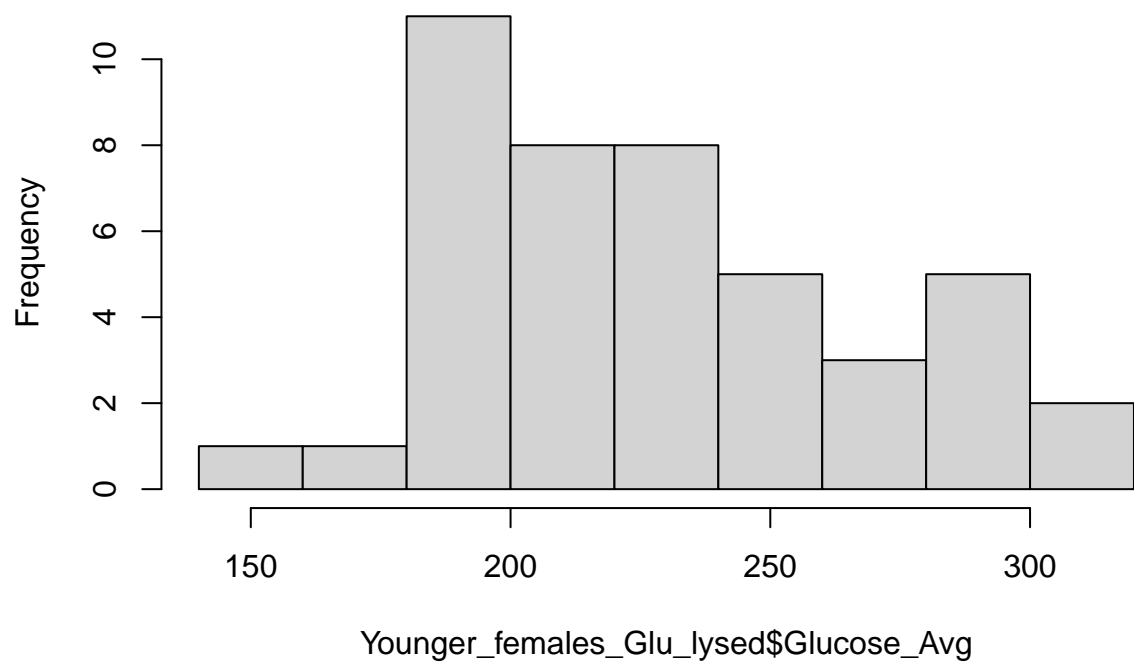
```
## [1] 157
```

```
max(Younger_females_Glu_lysed$Glucose_Avg)
```

```
## [1] 318.5
```

```
hist(Younger_females_Glu_lysed$Glucose_Avg)
```

Histogram of Younger_females_Glu_lysed$Glucose_Avg

Calcualte Sample size For Glucose Model

```
Younger_females_Glu_lysed %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##   Total Females
## 1    44      44
```

GLMER -> Glucose Model (Includes Lysed Plasma Individuals)

```
glmer_model_Glu <- glmer(Died_InTrt ~  Glucose_Avg +
                           (1 | Mom_ID) + (1 | Cohort),
                         data = Younger_females_Glu_lysed,
                         family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Glu)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Glucose_Avg + (1 | Mom_ID) + (1 | Cohort)
##    Data: Younger_females_Glu_lysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     51.2     58.2    -21.6      43.2        39
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7136 -0.5690 -0.4787 -0.3416  2.6328
##
## Random effects:
##  Groups Name        Variance  Std.Dev.
##  Mom_ID (Intercept) 1.135e-14 1.065e-07
##  Cohort (Intercept) 0.000e+00 0.000e+00
## Number of obs: 43, groups:  Mom_ID, 26; Cohort, 10
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.853865   2.303172   0.371    0.711
## Glucose_Avg -0.009738   0.010333  -0.942    0.346
##
## Correlation of Fixed Effects:
```

```
##              (Intr)
## Glucose_Avg -0.986
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

#Not significant p=0.346

Remove rows with plasma color other than ("U" or "U/L", "Y") and calculate how many individuals were removed

Logic: Hemolysis could release compounds (NADH, Catalase) that may interfere and skew glucose and ketone readings

```
# Remove rows with missing values
Younger_females_Glu_NotLysed <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Glucose_Avg))

# Capture dataset before plasma-color filtering
Younger_females_Glu_before_plasma <- Younger_females_Glu_NotLysed

# Remove individuals with lysed plasma appearance
Younger_females_Glu_NotLysed <- Younger_females_Glu_NotLysed %>%
  filter(Plasma_Color %in% c("U", "U/L", "Y"))

#Counts how many individuals were removed due to lysis or being yolky
removed_plasma_rows <- Younger_females_Glu_before_plasma %>%
  filter(!Plasma_Color %in% c("U", "U/L", "Y"))

# Total rows (samples) removed
n_removed_rows <- nrow(removed_plasma_rows)
cat("Rows removed due to plasma appearance:", n_removed_rows, "\n")
```

```
## Rows removed due to plasma appearance: 13
```

```
# Unique individuals removed (by ID_Band)
n_removed_individuals <- removed_plasma_rows %>%
  distinct(ID_Band) %>%
  nrow()
cat("Unique individuals removed:", n_removed_individuals, "\n")
```

```
## Unique individuals removed: 13
```

```
# 13 individuals removed due to plasma appearance

# Make a table of filtered out lysed individuals by Died_InTrt
removed_by_died_status <- removed_plasma_rows %>%
  distinct(ID_Band, Died_InTrt) %>%
  group_by(Died_InTrt) %>%
```

```
    summarize(n_removed = n(), .groups = "drop")

print(removed_by_died_status)
```

```
## # A tibble: 2 x 2
##   Died_InTrt n_removed
##   <fct>          <int>
## 1 No                 9
## 2 Yes                4
```

```
# 13 individuals removed due to plasma appearance, 9 that did not die, 4 that died
```

Recalculate Sample Size

```
Younger_females_Glu_NotLysed %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##   Total Females
## 1    31      31
```

Glucose Model (Excludes individuals with lysed plasma)

```
glmer_model_Glu <- glmer(Died_InTrt ~  Glucose_Avg +
                           (1 | Mom_ID) + (1 | Cohort),
                         data = Younger_females_Glu_NotLysed,
                         family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Glu)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Glucose_Avg + (1 | Mom_ID) + (1 | Cohort)
##    Data: Younger_females_Glu_NotLysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     31.2     36.8    -11.6     23.2        26
##
## Scaled residuals:
```

```
##      Min      1Q   Median      3Q      Max
## -1.04878 -0.02221 -0.01257 -0.00964  1.85115
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Mom_ID (Intercept) 128.2    11.32
##  Cohort (Intercept)   0.0     0.00
## Number of obs: 30, groups:  Mom_ID, 22; Cohort, 9
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -13.94126    8.97713  -1.553     0.12
## Glucose_Avg   0.02442    0.03096   0.789     0.43
##
## Correlation of Fixed Effects:
##            (Intr)
## Glucose_Avg -0.940
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Not significant p=0.43, high variance in Mom_ID. No cohort effect observed.
```

GLMER -> Ketone Model (Includes Lysed Plasma Individuals) -> To see whether baseline ketone levels predict death in younger females due to acute heat

Remove rows with missing values

```
Younger_females_Ket_Lysed <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Ketone_Avg))
```

Calculate Mean Ketone Values & plot residual

```
mean(Younger_females_Ket_Lysed$Ketone_Avg)
```

```
## [1] 0.5046512
```

```
sd(Younger_females_Ket_Lysed$Ketone_Avg)
```
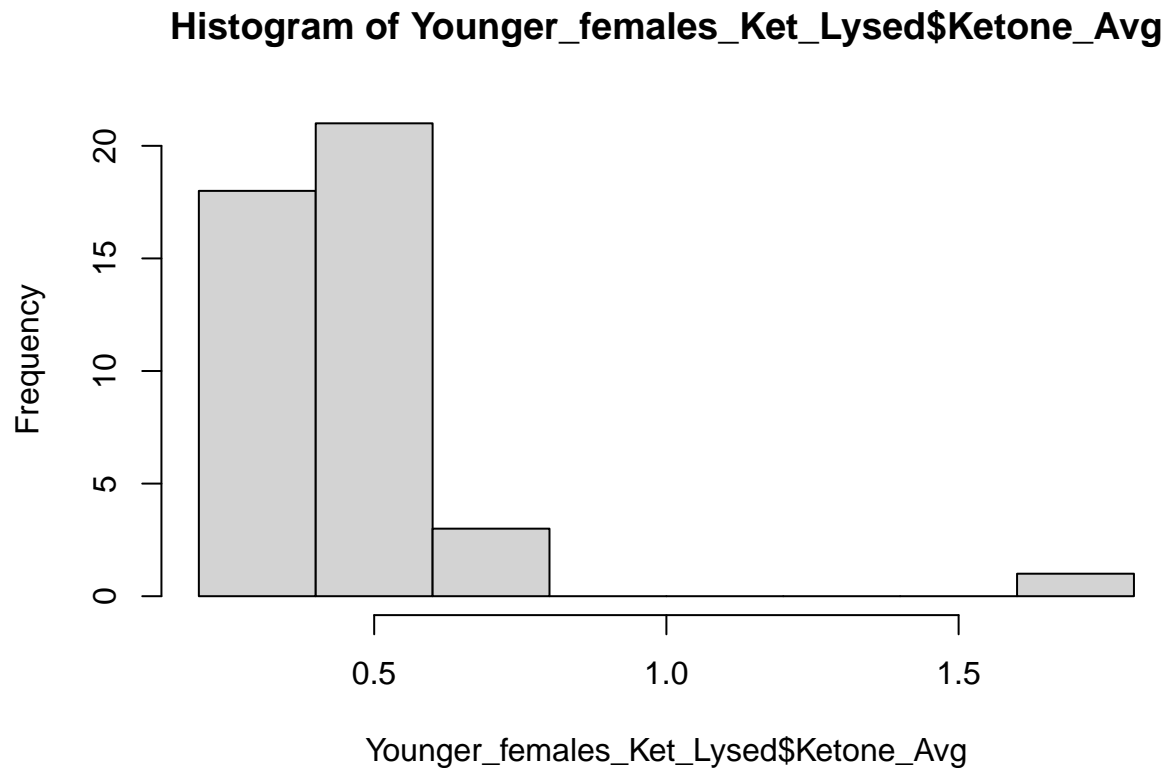
```
## [1] 0.2298586
```

```
min(Younger_females_Ket_Lysed$Ketone_Avg)
```

```
## [1] 0.3
```

```
max(Younger_females_Ket_Lysed$Ketone_Avg)
```

```
## [1] 1.8
```

```
hist(Younger_females_Ket_Lysed$Ketone_Avg)
```

Histogram of Younger_females_Ket_Lysed$Ketone_Avg

Calcualte Sample size For Ketone Model (Includes Lysed Plasma Appearance Individuals)

```
Younger_females_Ket_Lysed %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##   Total Females
## 1    43      43
```

GLMER -> Ketone Model (Includes Lysed Plasma Individuals)

```
glmer_model_Ket <- glmer(Died_InTrt ~  Ketone_Avg +
                          (1 | Mom_ID) + (1 | Cohort),
                        data = Younger_females_Ket_Lysed,
                        family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Ket)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Ketone_Avg + (1 | Mom_ID) + (1 | Cohort)
##    Data: Younger_females_Ket_Lysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     51.4     58.3    -21.7      43.4        38
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.5799 -0.5496 -0.5208 -0.4677  2.0263
##
## Random effects:
##  Groups Name         Variance  Std.Dev.
##  Mom_ID (Intercept) 9.733e-15 9.866e-08
##  Cohort (Intercept) 1.918e-16 1.385e-08
## Number of obs: 42, groups:  Mom_ID, 26; Cohort, 10
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.7669     1.1930  -0.643    0.520
## Ketone_Avg   -1.0758     2.3471  -0.458    0.647
##
## Correlation of Fixed Effects:
##            (Intr)
## Ketone_Avg -0.949
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

Remove rows with plasma color other than ("U" or "U/L", "Y") and calculate how many individuals were removed

Logic: Hemolysis could release compounds (NADH, Catalase) that may interfere and skew glucose and ketone readings

```r
# Remove rows with missing values
Younger_females_Ket_NotLysed <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Ketone_Avg))

# Capture dataset before plasma-color filtering
Younger_females_Ket_before_plasma <- Younger_females_Ket_NotLysed

# Remove individuals with lysed plasma appearance
Younger_females_Ket_NotLysed <- Younger_females_Ket_NotLysed %>%
  filter(Plasma_Color %in% c("U", "U/L", "Y"))

#Counts how many individuals were removed due to lysis or being yolky
removed_plasma_rows_ket <- Younger_females_Ket_before_plasma %>%
  filter(!Plasma_Color %in% c("U", "U/L", "Y"))

# Total rows (samples) removed
n_removed_rows_ket <- nrow(removed_plasma_rows_ket)
cat("Rows removed due to plasma appearance:", n_removed_rows_ket, "\n")
```

```
## Rows removed due to plasma appearance: 12
```

```r
# Unique individuals removed (by ID_Band)
n_removed_individuals_ket <- removed_plasma_rows_ket %>%
  distinct(ID_Band) %>%
  nrow()
cat("Unique individuals removed:", n_removed_individuals_ket, "\n")
```

```
## Unique individuals removed: 12
```

```r
# 12 individuals removed due to plasma appearance
```

Recalculate Sample Size

```r
Younger_females_Ket_NotLysed %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##    Total Females
## 1    31       31
```

```
# Sample Size: 31
```

Ketone Model (Excludes individuals with lysed plasma)

```
glmer_model_Ket <- glmer(Died_InTrt ~  Ketone_Avg +
                           (1 | Mom_ID) + (1 | Cohort),
                         data = Younger_females_Ket_NotLysed,
                         family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Ket)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Ketone_Avg + (1 | Mom_ID) + (1 | Cohort)
##    Data: Younger_females_Ket_NotLysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     31.9     37.5    -11.9      23.9        26
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.92984 -0.01986 -0.01857 -0.01793  1.46052
##
## Random effects:
##  Groups Name         Variance Std.Dev.
##  Mom_ID (Intercept) 107.3    10.36
##  Cohort (Intercept)   0.0     0.00
## Number of obs: 30, groups:  Mom_ID, 22; Cohort, 9
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.22423    0.00416 -1977.0   <2e-16 ***
## Ketone_Avg   0.72241    0.00416   173.7   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr)
## Ketone_Avg 0.000
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# Ketones is significant p=<2e-16 ***, high variance in Mom_ID. No cohort effect observed.
```

Hematocrit Model ## To see whether baseline hematocrit levels predict death in younger females due to acute heat.

Remove rows with missing values

```
Younger_females_Hem <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Hematocrit))
```

Calculate Mean Hematocrit Values & plot residuals

```
mean(Younger_females_Hem$Hematocrit)
```

```
## [1] 53.9875
```

```
sd(Younger_females_Hem$Hematocrit)
```
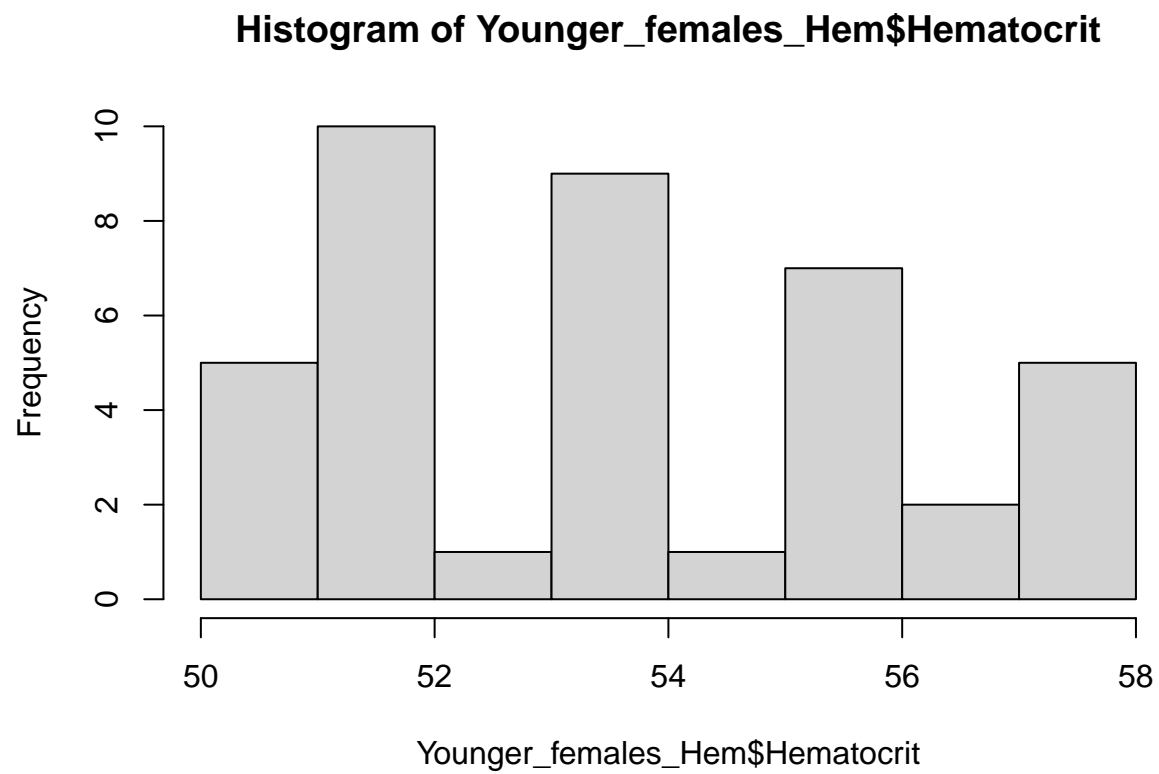
```
## [1] 2.507649
```

```
min(Younger_females_Hem$Hematocrit)
```

```
## [1] 50
```

```
max(Younger_females_Hem$Hematocrit)
```

```
## [1] 58
```
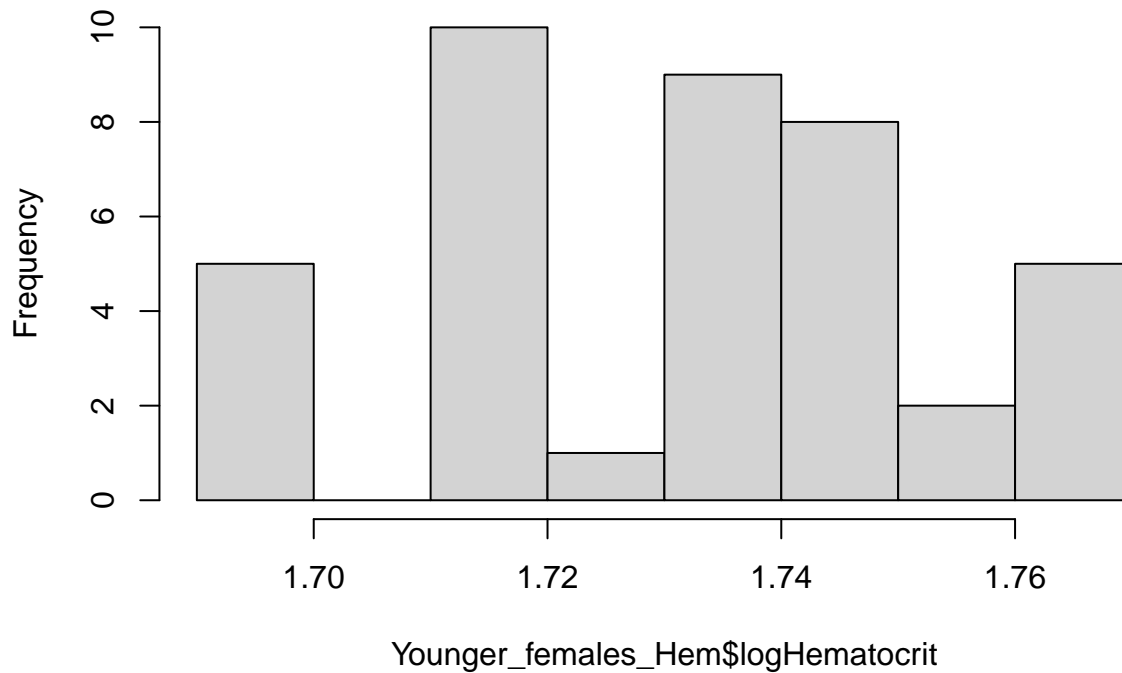
```
hist(Younger_females_Hem$Hematocrit)
```

Histogram of Younger_females_Hem$Hematocrit

#Create log-transformed Hematocrit columns

```
Younger_females_Hem$logHematocrit <- log10(Younger_females_Hem$Hematocrit)
hist(Younger_females_Hem$logHematocrit)
```

Histogram of Younger_females_Hem$logHematocrit

Younger_females_Hem$logHematocrit

```
#Better
```

Calculate Sample Size

```
Younger_females_Hem %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##   Total Females
## 1    40      40
```

```
# Sample Size: 40
```

GLMER -> Hematocrit Model (Includes Lysed Individuals)

```r
glmer_model_Hematocrit <- glmer(Died_InTrt ~  logHematocrit +
                                  (1 | Mom_ID) + (1 | Cohort),
                                data = Younger_females_Hem,
                                family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
summary(glmer_model_Hematocrit)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ logHematocrit + (1 | Mom_ID) + (1 | Cohort)
##     Data: Younger_females_Hem
##
##       AIC       BIC    logLik -2*log(L)  df.resid
##      47.0      53.6     -19.5      39.0        35
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.6093 -0.4937 -0.4094 -0.3622  2.1411
##
## Random effects:
##  Groups Name         Variance Std.Dev.
##  Mom_ID (Intercept) 0.3747   0.6121
##  Cohort (Intercept) 0.0000   0.0000
## Number of obs: 39, groups:  Mom_ID, 26; Cohort, 9
##
## Fixed effects:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -27.44       37.05  -0.741    0.459
## logHematocrit   14.98       21.36   0.701    0.483
##
## Correlation of Fixed Effects:
##            (Intr)
## logHematcrt -1.000
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# Hematocrit not significant p=0.483 - low variance in mom ID, no cohort effect observed
```

Remove rows with plasma color other than ("U" or "U/L", "Y") and calculate how many individuals were removed

Logic: Hemolysis could change levels of hematocrit measured specifically by decreasing the packed RBC's skewing data towards higher Hct

```
# Remove rows with missing values
Younger_females_Hem_NotLysed <- AHDB_Baseline_YoungerFemale %>%
  filter(!is.na(Hematocrit))

# Capture dataset before plasma-color filtering
Younger_females_Hem_before_plasma <- Younger_females_Hem_NotLysed

# Remove individuals with lysed plasma appearance
Younger_females_Hem_NotLysed <- Younger_females_Hem_NotLysed %>%
  filter(Plasma_Color %in% c("U", "U/L", "Y"))

#Counts how many individuals were removed due to lysis or being yolky
removed_plasma_rows <- Younger_females_Hem_before_plasma %>%
  filter(!Plasma_Color %in% c("U", "U/L", "Y"))

# Total rows (samples) removed
n_removed_rows <- nrow(removed_plasma_rows)
cat("Rows removed due to plasma appearance:", n_removed_rows, "\n")
```

```
## Rows removed due to plasma appearance: 10
```

```
# Unique individuals removed (by ID_Band)
n_removed_individuals <- removed_plasma_rows %>%
  distinct(ID_Band) %>%
  nrow()
cat("Unique individuals removed:", n_removed_individuals, "\n")
```

```
## Unique individuals removed: 10
```

```
# 10 individuals removed due to plasma appearance
```

Recalculate Sample Size

```
Younger_females_Hem_NotLysed %>%
  distinct(ID_Band, Sex) %>%
  summarise(
    Total   = n(),
    Females = sum(Sex == "F")
  )
```

```
##   Total Females
## 1    30       30
```

#Create log-transformed Hematocrit columns

```
Younger_females_Hem_NotLysed$logHematocrit <- log10(Younger_females_Hem_NotLysed$Hematocrit)
hist(Younger_females_Hem_NotLysed$logHematocrit)
```

Histogram of Younger_females_Hem_NotLysed$logHematocrit

Younger_females_Hem_NotLysed$logHematocrit

GLMER -> Hematocrit Model (Excludes Lysed Individuals)

```
glmer_model_Hematocrit <- glmer(Died_InTrt ~  logHematocrit +
                            (1 | Mom_ID) + (1 | Cohort),
                        data = Younger_females_Hem_NotLysed,
                        family = binomial)

summary(glmer_model_Hematocrit)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
## Family: binomial  ( logit )
```

```
## Formula: Died_InTrt ~ logHematocrit + (1 | Mom_ID) + (1 | Cohort)
##    Data: Younger_females_Hem_NotLysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     32.3     37.8    -12.2     24.3        25
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.6081 -0.4078 -0.2636 -0.1759  2.1689
##
## Random effects:
##  Groups Name          Variance Std.Dev.
##  Mom_ID (Intercept) 0.8062   0.8979
##  Cohort (Intercept) 0.6247   0.7904
## Number of obs: 29, groups:  Mom_ID, 21; Cohort, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -73.74      61.58  -1.198    0.231
## logHematocrit   41.27      35.25   1.171    0.242
##
## Correlation of Fixed Effects:
##            (Intr)
## logHematcrt -1.000

# Not significant p=0.242, low variance for cohort and mom_id
```

ALL SUBSEQUENT ANALYSES NOW REMOVE MOM_ID AS A RANDOM EFFECT

GLMER -> Glucose Model (WITH LYSED BUT NO Mom_ID) **Brynleigh ESEB Poster

```
glmer_model_Glu <- glmer(Died_InTrt ~  Glucose_Avg +
                           (1 | Cohort),
                       data = Younger_females_Glu_lysed,
                       family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```
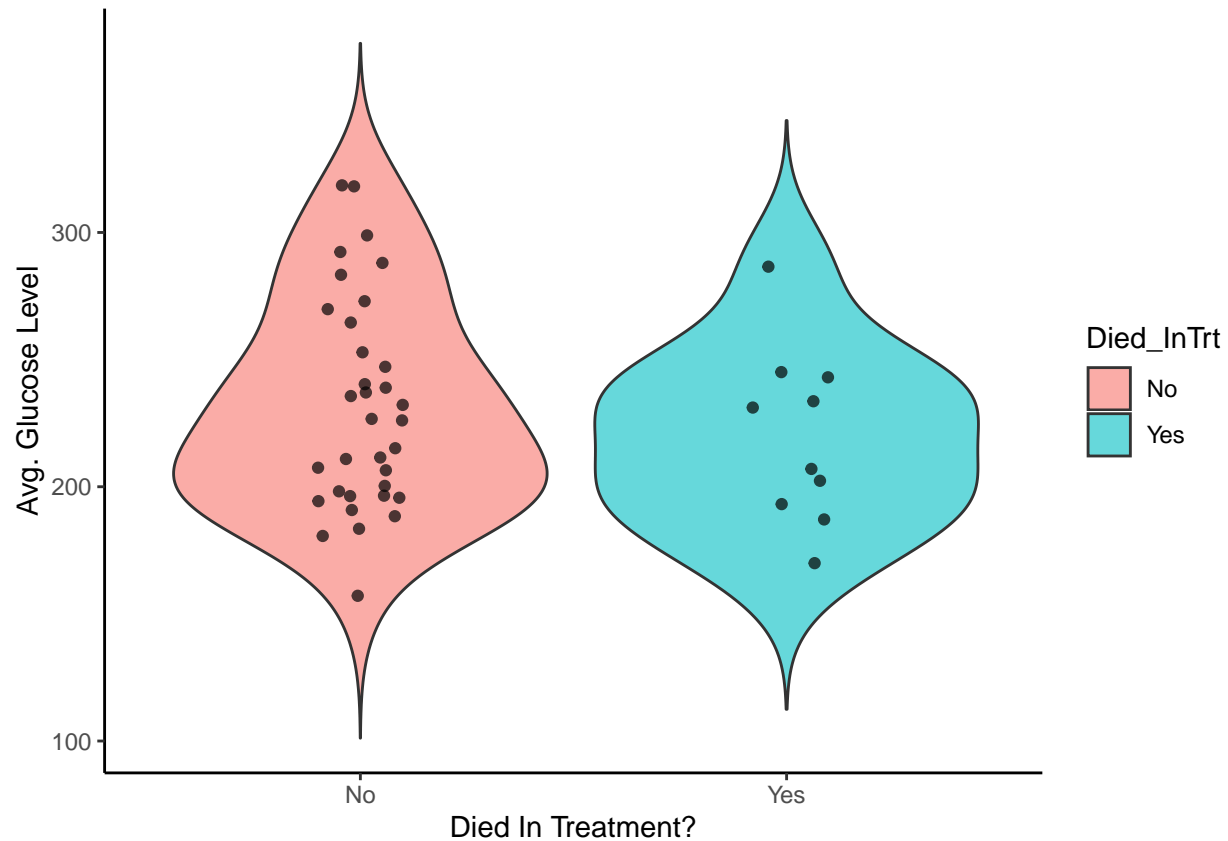
```
summary(glmer_model_Glu)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Glucose_Avg + (1 | Cohort)
##    Data: Younger_females_Glu_lysed
##
```

```
##       AIC      BIC    logLik -2*log(L)  df.resid
##      52.5     57.8     -23.2      46.5        41
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7125 -0.5886 -0.5126 -0.3752  2.3499
##
## Random effects:
##  Groups Name         Variance Std.Dev.
##  Cohort (Intercept) 0        0
## Number of obs: 44, groups:  Cohort, 10
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.571871   2.191323   0.261    0.794
## Glucose_Avg -0.007960   0.009725  -0.819    0.413
##
## Correlation of Fixed Effects:
##             (Intr)
## Glucose_Avg -0.986
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# No change in results (p=0.413)

# Plot
ggplot(Younger_females_Glu_lysed, aes(x = Died_InTrt, y = Glucose_Avg)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.6) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  labs(
    x = "Died In Treatment?",
    y = "Avg. Glucose Level"
  ) +
  theme_classic()
```

GLMER -> Glucose Model (WITHOUT LYSED INDIVIDUALS & Mom_ID)

```
glmer_model_Glu <- glmer(Died_InTrt ~  Glucose_Avg +
                            (1 | Cohort),
                         data = Younger_females_Glu_NotLysed,
                         family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?
```

```
summary(glmer_model_Glu)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Glucose_Avg + (1 | Cohort)
##    Data: Younger_females_Glu_NotLysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     35.6     39.9    -14.8      29.6        28
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7349 -0.4318 -0.3156 -0.2833  1.9116
##
```

```
## Random effects:
##  Groups Name           Variance Std.Dev.
##  Cohort (Intercept) 1.493     1.222
## Number of obs: 31, groups:  Cohort, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.615251   3.562556  -0.734    0.463
## Glucose_Avg  0.004064   0.015175   0.268    0.789
##
## Correlation of Fixed Effects:
##             (Intr)
## Glucose_Avg -0.973
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```
# No change in results (p=0.789)
```

GLMER -> Ketone Model (WITH LYSED BUT NO Mom_ID) **Brynleigh ESEB Poster

```
glmer_model_Ket <- glmer(Died_InTrt ~  Ketone_Avg +
                             (1 | Cohort),
                         data = Younger_females_Ket_Lysed,
                         family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Ket)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Ketone_Avg + (1 | Cohort)
##    Data: Younger_females_Ket_Lysed
##
##      AIC       BIC    logLik -2*log(L)  df.resid
##     52.0      57.2     -23.0      46.0        40
##
## Scaled residuals:
##    Min      1Q  Median      3Q     Max
## -0.6550 -0.5955 -0.5413 -0.3020  2.0323
##
## Random effects:
##  Groups Name          Variance  Std.Dev.
##  Cohort (Intercept) 3.431e-16 1.852e-08
## Number of obs: 43, groups:  Cohort, 10
##
```

```
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.2739      1.3502  -0.203     0.839
## Ketone_Avg   -1.9075      2.7901  -0.684     0.494
##
## Correlation of Fixed Effects:
##            (Intr)
## Ketone_Avg -0.963
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
# No change in results (p=0.494)

# Calculate z-scores for outliers
Younger_females_Ket_Lysed <- Younger_females_Ket_Lysed %>%
  mutate(z_score = (Ketone_Avg - mean(Ketone_Avg, na.rm = TRUE)) / sd(Ketone_Avg, na.rm = TRUE))
# Identify outliers with z-scores > 3 or < -3
outliers <- Younger_females_Ket_Lysed %>%
  filter(abs(z_score) > 3)
print(outliers)
```

```
##   AgeCategory ID_Band Sex Mom_ID Cohort Treatment Died_InTrt TimePoint
## 1     Younger    4703   F   4112      5         B         No  Baseline
##   Glucose_1 Glucose_2 Glucose_3 Ketone_1 Ketone_2 Ketone_3 Plasma_Color
## 1       178       183        NA      1.8      1.8       NA            U
##   Hematocrit BleedTime_Min Glucose_Meter_Lot_Number Ketone_Meter_Lot_Number
## 1         58          1:20                  042224A                 75001L04
##   Room Notes Glucose_Avg Ketone_Avg  z_score
## 1  118             180.5        1.8 5.635415
```

```
# Remove Outliers & rerun stats
Younger_females_Ket_Lysed_clean <- Younger_females_Ket_Lysed %>%
  filter(abs(z_score) <= 3)

glmer_model_Ket_No_Outliers <- glmer(Died_InTrt ~  Ketone_Avg +
                         (1 | Cohort),
                  data = Younger_females_Ket_Lysed_clean,
                  family = binomial)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(glmer_model_Ket_No_Outliers)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Ketone_Avg + (1 | Cohort)
##    Data: Younger_females_Ket_Lysed_clean
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     51.9     57.1    -22.9      45.9        39
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.6368 -0.5895 -0.5458 -0.4679  1.9788
##
## Random effects:
##  Groups Name        Variance  Std.Dev.
##  Cohort (Intercept) 5.903e-15 7.683e-08
## Number of obs: 42, groups:  Cohort, 10
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.4405     1.6071  -0.274    0.784
## Ketone_Avg   -1.5408     3.3725  -0.457    0.648
##
## Correlation of Fixed Effects:
##            (Intr)
## Ketone_Avg -0.974
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# No change in results (p-value=0.648)

# Plot
ggplot(Younger_females_Ket_Lysed_clean, aes(x = Died_InTrt, y = Ketone_Avg)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.6) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  labs(
    x = "Died In Treatment?",
    y = "Avg. Ketone Level"
  ) +
  theme_classic()
```
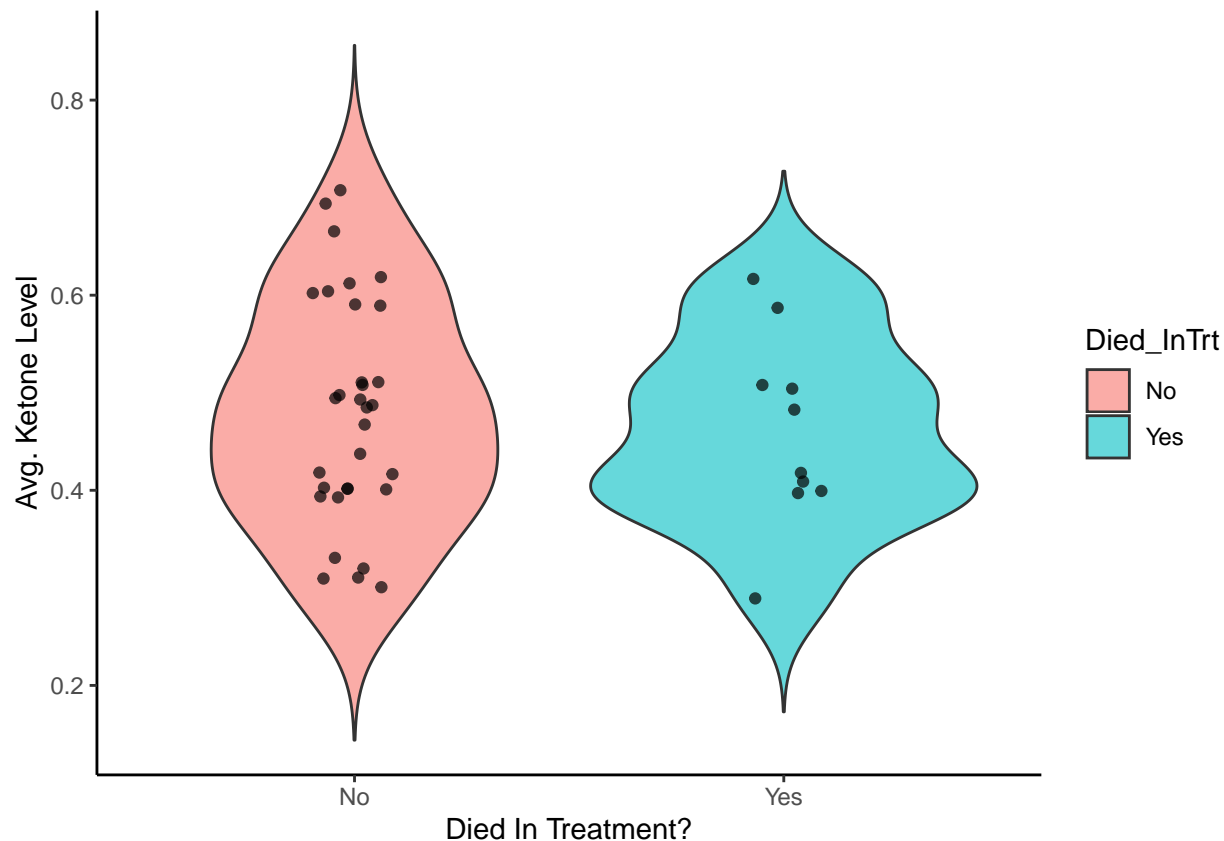
GLMER -> Ketone Model (WITHOUT LYSED INDIVIDUALS & Mom_ID)

```
glmer_model_Ket <- glmer(Died_InTrt ~ Ketone_Avg +
                               (1 | Cohort),
                         data = Younger_females_Ket_NotLysed,
                         family = binomial)

summary(glmer_model_Ket)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Ketone_Avg + (1 | Cohort)
##    Data: Younger_females_Ket_NotLysed
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##     35.6     39.9    -14.8      29.6        28
##
## Scaled residuals:
##    Min     1Q  Median     3Q    Max
## -0.6617 -0.4360 -0.3473 -0.2956  2.0528
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Cohort (Intercept) 1.035    1.017
## Number of obs: 31, groups:  Cohort, 9
```

```
## 
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.104      1.921  -0.575    0.566
## Ketone_Avg    -1.108      3.462  -0.320    0.749
## 
## Correlation of Fixed Effects:
##            (Intr)
## Ketone_Avg -0.918
```

```
# Ketones are no longer significant (p-value = 0.749)
```

GLMER -> Hematocrit Model (WITH LYSED BUT NO Mom_ID)

```
glmer_model_Hem <- glmer(Died_InTrt ~  logHematocrit +
                           (1 | Cohort),
                       data = Younger_females_Hem,
                       family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0096257 (tol = 0.002, component 1)
```

```
summary(glmer_model_Hem)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ logHematocrit + (1 | Cohort)
##    Data: Younger_females_Hem
## 
##      AIC      BIC   logLik -2*log(L)  df.resid
##     47.9     53.0    -21.0      41.9        37
## 
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7187 -0.5442 -0.4618 -0.3246  2.0627
## 
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Cohort (Intercept) 0.4044   0.6359
## Number of obs: 40, groups:  Cohort, 9
## 
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -28.247922   0.007375   -3830   <2e-16 ***
## logHematocrit 15.562166   0.007379    2109   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Correlation of Fixed Effects:
##            (Intr)
## logHematcrt 0.001
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.0096257 (tol = 0.002, component 1)
```

```r
# Convergence Error at Maximum

glmer_model_Hem <- glmer(Died_InTrt ~  Hematocrit +
                                (1 | Cohort),
                         data = Younger_females_Hem,
                         family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidenti
##  - Rescale variables?
```

```r
summary(glmer_model_Hem)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Hematocrit + (1 | Cohort)
##     Data: Younger_females_Hem
## 
##       AIC      BIC   logLik -2*log(L)  df.resid
##        48       53      -21        42        37
## 
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7243 -0.5427 -0.4611 -0.3224  2.0747
## 
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Cohort (Intercept) 0.419    0.6473
## Number of obs: 40, groups:  Cohort, 9
## 
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -8.2196     9.3861  -0.876    0.381
## Hematocrit    0.1282     0.1719   0.746    0.456
## 
## Correlation of Fixed Effects:
##            (Intr)
## Hematocrit -0.999
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```r
# Scaling Error

# Scale Predictor
```

```r
Younger_females_Hem <- Younger_females_Hem %>%
  mutate(
    Hematocrit_s = as.numeric(scale(Hematocrit))
  )

glmer_model_Hem <- glmer(Died_InTrt ~  Hematocrit_s +
                           (1 | Cohort),
                         data = Younger_females_Hem,
                         family = binomial)

summary(glmer_model_Hem)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ Hematocrit_s + (1 | Cohort)
##    Data: Younger_females_Hem
##
##      AIC      BIC   logLik -2*log(L)  df.resid
##       48       53      -21        42        37
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7243 -0.5427 -0.4611 -0.3224  2.0747
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Cohort (Intercept) 0.419    0.6473
## Number of obs: 40, groups:  Cohort, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.2967     0.4806  -2.698  0.00697 **
## Hematocrit_s   0.3216     0.4312   0.746  0.45578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Hematocrt_s -0.192
```
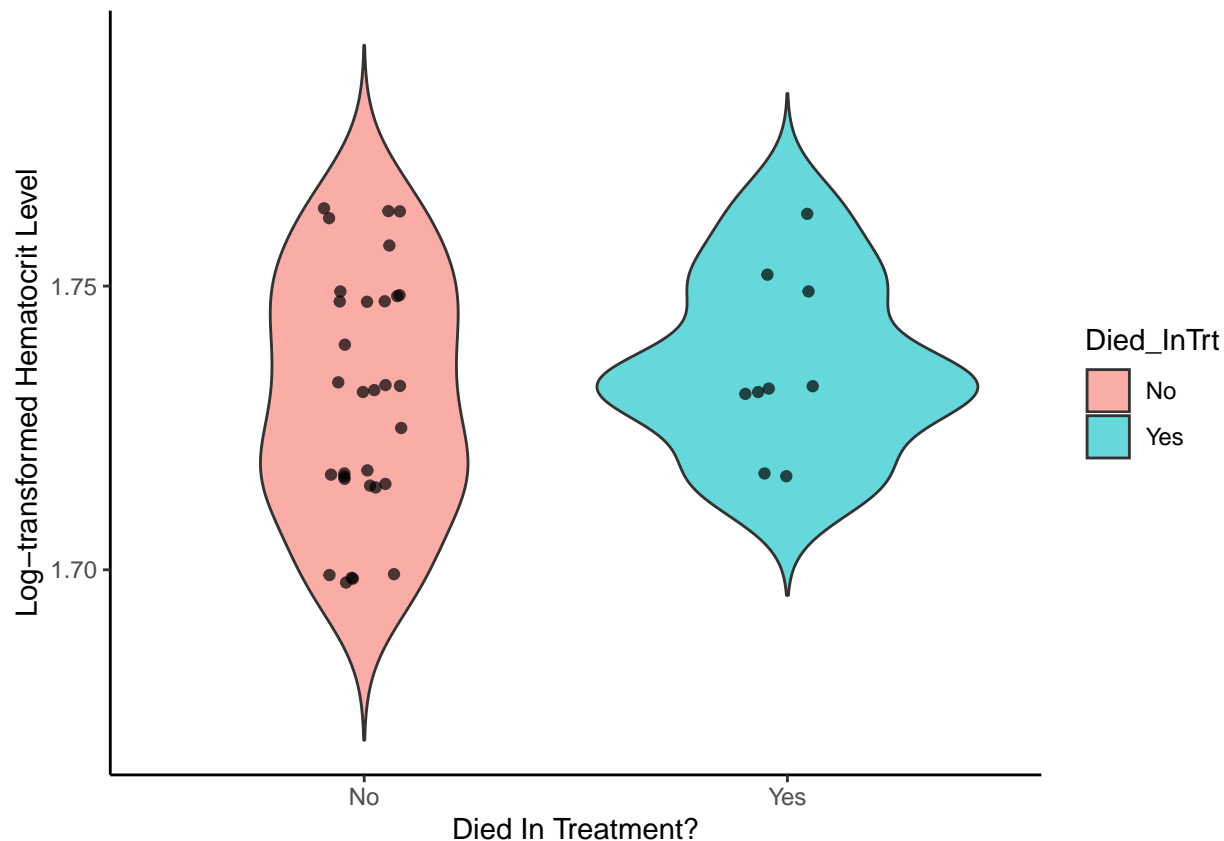
```r
# Results are same as before i.e. hematocrit is not significant

# Plot
ggplot(Younger_females_Hem, aes(x = Died_InTrt, y = logHematocrit)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.6) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  labs(
    x = "Died In Treatment?",
    y = "Log-transformed Hematocrit Level"
  ) +
  theme_classic()
```

GLMER -> Hematocrit Model (WITHOUT LYSED INDIVIDUALS & Mom_ID) **Brynleigh ESEB Poster

```
glmer_model_Hem <- glmer(Died_InTrt ~  logHematocrit +
                         (1 | Cohort),
                    data = Younger_females_Hem_NotLysed,
                    family = binomial)

summary(glmer_model_Hem)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Died_InTrt ~ logHematocrit + (1 | Cohort)
##    Data: Younger_females_Hem_NotLysed
##
##      AIC      BIC    logLik -2*log(L)  df.resid
##     32.6     36.8     -13.3      26.6        27
##
## Scaled residuals:
##     Min       1Q   Median       3Q      Max
## -0.86453 -0.38090 -0.25070 -0.08391  2.41342
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Cohort (Intercept) 3.461    1.86
```

43

```
## Number of obs: 30, groups:   Cohort, 9
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -85.89      68.70   -1.25    0.211
## logHematocrit   48.29      39.24    1.23    0.219
##
## Correlation of Fixed Effects:
##              (Intr)
## logHematcrt -1.000
```

```
# Same results as before, hematocrit is not significant (p=0.219)

# Plot
ggplot(Younger_females_Hem_NotLysed, aes(x = Died_InTrt, y = logHematocrit)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.6) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  labs(
    x = "Died In Treatment?",
    y = "Log-transformed Hematocrit Level"
  ) +
  theme_classic()
```
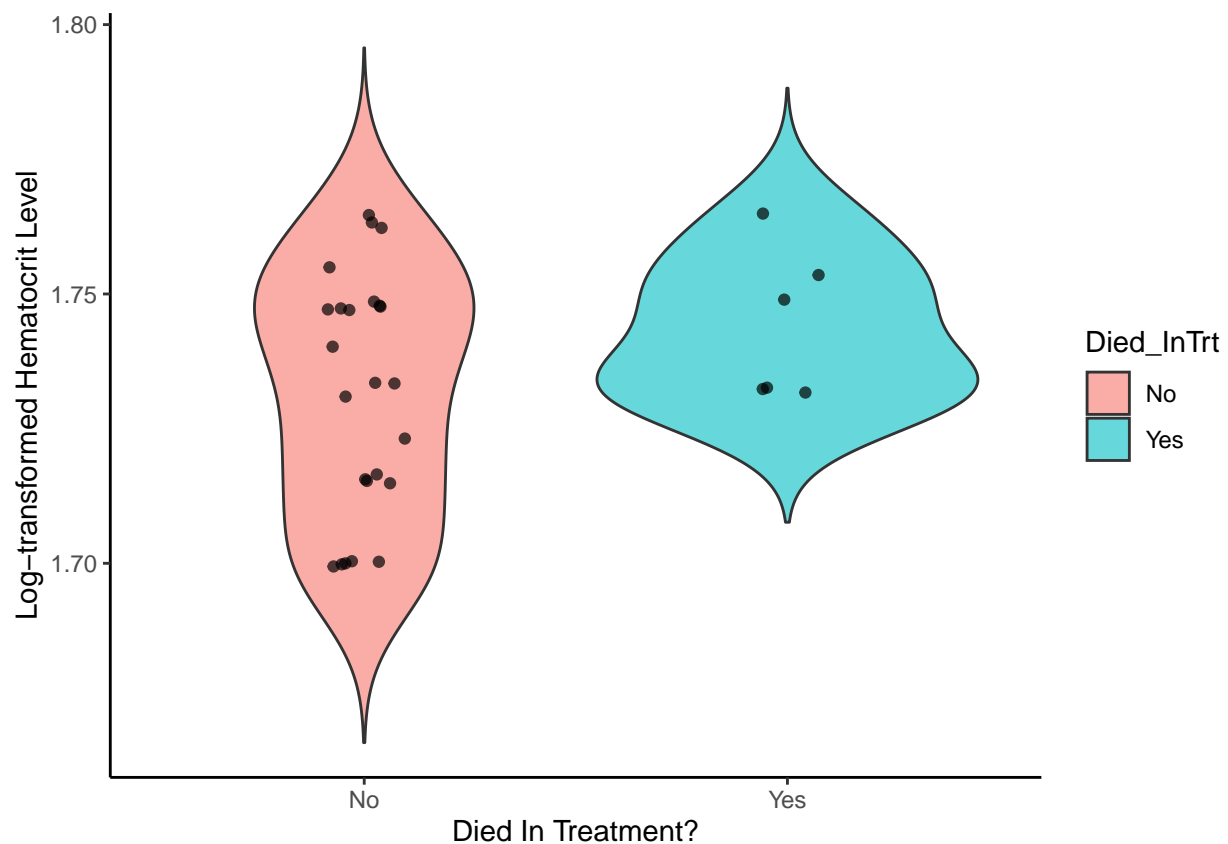
ESEB Poster Brynleigh -Graph of Glucose, Ketone, Hematocrit averages

Models do not include mom ID due to only two being the same within the died in treatment individuals

Includes cohort as a random effect

Glucose/Ketone includes lysed individuals because the significance did not change for any model with/without , Hematocrit does not include lysed because that is when we record the lyses occuring

```r
# Glucose Plot
p1 <- ggplot(Younger_females_Glu_lysed, aes(x = Died_InTrt, y = Glucose_Avg)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.9) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  scale_fill_manual(values = c("#C0504D", "#153F70")) +  # Custom colors
  labs(
    x = "Died In Treatment?",
    y = "Avg. Glucose"
  ) +
  theme_classic(base_size = 18) +
  theme(legend.position = "none")  # Removes the legend

# Ketone Plot
p2 <- ggplot(Younger_females_Ket_Lysed_clean, aes(x = Died_InTrt, y = Ketone_Avg)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.9) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  scale_fill_manual(values = c("#C0504D", "#153F70")) +  # Custom colors
  labs(
    x = "Died In Treatment?",
    y = "Avg. Ketone"
  ) +
  theme_classic(base_size = 18) +
  theme(legend.position = "none")  # Removes the legend

# Hematocrit Plot
p3 <- ggplot(Younger_females_Hem_NotLysed, aes(x = Died_InTrt, y = logHematocrit)) +
  geom_violin(aes(fill = Died_InTrt), trim = FALSE, alpha = 0.9) +  # Violin by group
  geom_jitter(width = 0.1, alpha = 0.7, size = 1.5, color = "black") +  # Add data points
  scale_fill_manual(values = c("#C0504D", "#153F70")) +  # Custom colors
  labs(
    x = "Died In Treatment?",
    y = "Log-transformed Hematocrit"
  ) +
  theme_classic(base_size = 18) +
  theme(legend.position = "none")  # Removes the legend

library(cowplot)
```
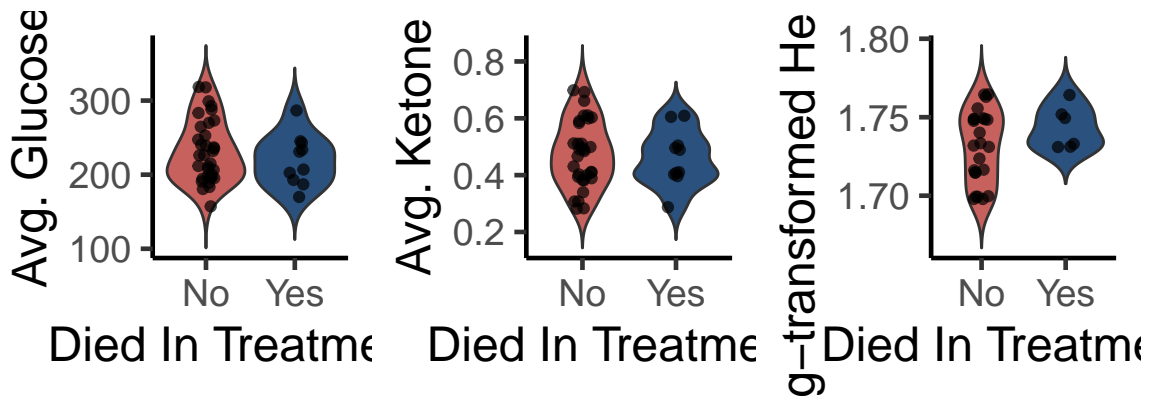
```
##
## Attaching package: 'cowplot'

## The following object is masked from 'package:ggeffects':
##
##     get_title
```

```
# Combine into 1 row and 3 columns
combined_blood_measure_plot <- plot_grid(p1, p2, p3, nrow = 1)
combined_blood_measure_plot
```



```
#ggsave("Combined_Blood_Plot.png", combined_blood_measure_plot, width = 12, height = 4)
```